Faculty and Researcher Publications

2008-04

Team 6: Enhanced Design of Experiment for Testing in a Joint Environment

Beach, Timothy

http://hdl.handle.net/10945/35640

Team 6: Enhanced Design of Experiment for Testing in a Joint Environment

TEAM 6 MEMBERS

Timothy BEACH – Lead, Contact
Col Eileen BJORKMAN
Dave DRYER, Ph.D.
Mark FIEBRANDT
Maureen SHORT
Al SCIARRETTA
Joint Test and Evaluation Methodology, USA

David KELTON, Ph.D.
University of Cincinnati, USA

Susan SANCHEZ, Ph.D.
Steve UPTON
Naval Postgraduate School (NPS), USA

LTC (P) Jeff SHAMBURG, Ph.D.
MAJ Eric TOLLEFSON
MAJ Jonathan ALT
TRADOC Analysis Center (TRAC), USA

Thomas DONNELLY, Ph.D.
US Army Edgewood CB Center, USA

Lawton CLITES
Digital Consulting Services, USA

INTRODUCTION - TEAM 6

The intent of Team 6 activities at the International Data Farming Workshop (IDFW) 16 was to explore enhanced design of experiment (DOE) techniques and models relevant to developing evaluation strategies for testing in a joint environment (TIJE). This goal was met through the utilization of the Map Aware Non-uniform Automata (MANA) model to trace a "call for fire" (CFF) from the originator to the final weapon system, at the detailed level of an individual task thread. A capability-level evaluation strategy for battlespace deconfliction tasks was used as the scenario driver for the data farming runs. This evaluation strategy has been developed as part of the Joint Test and Evaluation Methodology (JTEM) project.

JTEM is developing and enhancing the Capability Test Methodology (CTM) as best practice methods and processes for designing and executing testing in a joint environment. Part of the problem space JTEM has discovered in developing this methodology is when moving from system to system of systems (SoS), or to capability-focused test and evaluation, the number of factors that are part of the test space grows significantly, even exponentially. Thus, part of the JTEM project is the mission to develop processes for refining this test space, based on DOE techniques for large factor; multiple response designs.

The planning, execution, and analysis of Team 6's data farming activities were completed within the context of the CTM's Develop Evaluation Strategy process. This process includes efficient DOEs, the use of computing clusters, and the iterative data farming process. Questions JTEM specifically wanted to focus on during IDFW 16 were:

- Given a critical joint issue (CJI) for battlespace deconfliction, which factors are the most important to examine for testing?
- What are some appropriate design of experiment techniques that could be applied to the test space?
- What data exploration and analysis methods would be appropriate to apply with so many factors?

Scenario

Prior to the actual execution of the workshop, Team 6 began to develop the use-case scenario shown in Figure 1. This scenario focuses on a joint forcible entry operation where friendly forces would be conducting joint fires, joint close air support, and close combat attack operations. These operations would expand a Blue (friendly) force lodgment and allow for control of key infrastructure in order to facilitate rapid force build-up in the joint operations area (JOA).

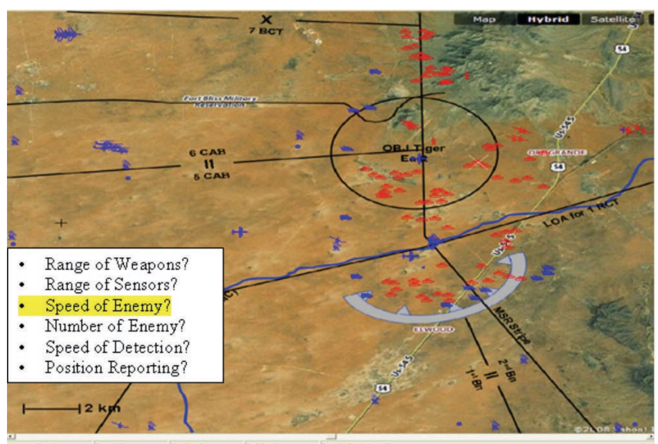

Figure 1: Team 6 Scenario

The developed mission desired effect was for threat forces to be destroyed or neutralized in the JOA. Once this scenario was developed, JTEM wanted to analyze different DOE techniques considered to be best practice, as well as look at promising new DOE techniques under development. The goal was to enhance the CTM methodology and incorporate the most current practices being applied in both industry and government laboratories. A screenshot of the MANA scenario shows the Blue (friendly) and threat forces with a list of potential influential factors. The general approach of Team 6 during IDFW 16 was to examine the developed scenario, apply appropriate DOE techniques, run thousands of iterations based on the applied DOE, and then analyze the results of the runs. By exercising this methodology, the team hoped to show data farming and DOE applications are extremely useful tools for test planning.

TIJE processes must develop critical evaluation issues to assess performance as it pertains to capabilities supporting joint missions. To address joint capability contributions to achieving desired mission effects, JTEM has developed the concept of a CJI. The CJI for a test should address the achievement of mission desired effects, the SoS' ability to accomplish joint operational tasks, and/or the SoS, system, or service attribute performance. The essential elements of a CJI include a capability's essential tasks, mission desired effects, Blue SoS aspects (across Doctrine, Organization, Training, Materiel, Leadership and Education, Personnel and Facilities--DOTMLPF), and conditions involving threat and environmental factors. These essential elements are contained in the capability crosswalk. A portion of the capability crosswalk for Team 6 is captured in Figure 2. It is important to state how the test issue contributes to achieving the desired mission end state outcomes in terms of mission desired effects. The CJIs should address the SoS capability to perform joint operational tasks and/or the SoS, system, or service attribute performance. CJIs can be of assistance to the appropriate authority when deciding whether to allow the SoS to advance to the next phase of development.

# MANA	Factors	MANA	MANA Parameter	Parameter Level
	Airspace Control		N/A	
	Assess the ability to perform battlespace deconfliction by Current/Future C2 SoS under a full range of military operations in order to achieve a destroyed or neutralized threat forces in the Joint Operations Area.	**Describe how will model the respective factor within the specific simulation**		
1	Current TBMCS/TAIS: Information not available for other Services	Inorganic SA commo links can be set up between any entity groups, thus allowing you to model which services can share information and who knows what. Can vary Inorganic SA Latency, Message Capacity, Buffer Capacity, Contact Type Accuracy (for friendly, enemy, neutral), Contact Location Accuracy (distance) and Reliability (of the message getting through) in order to get at the difference in solutions.	Range of SA for forward observer (Intersquad)	0 pixels (23.7 meters per pixel)
	Future JASMAD/JAMUS: Information available on all aircraft regardless of Service	Same as above		150 pixels
2	Current TBMCS/TAIS: Information not available for other Services	FO Current	Inorganic SA Fuse Time (Intersquad)	300 seconds
	Future JASMAD/JAMUS: Information available on all aircraft regardless of Service	FO New		0 seconds
		JTA C Current	Inorganic SA Fuse Time (Intersquad)	300 seconds
		JTAC New		0 seconds
		UAS Current	Inorganic SA Fuse Time (Intersquad)	300 seconds
		UAS New		0 seconds
3	Current TBMCS/TAIS: Information not available for other Services	FO Current	Contact Location Accuracy	50%
	Future JASMAD/JAMUS: Information available on all aircraft regardless of Service	FO New		90%
		JTA C Current		50%
		JTAC New		90%
		UAS Current		50%
		UAS New		90%
4	Current TBMCS/TAIS: Information not available for other Services	FO Current	Self- Intersquad SA (time steps)	60
	Future JASMAD/JAMUS: Information available on all aircraft regardless of Service	FO New		5

Figure 2: Part of Capability Crosswalk Developed By Team 6

An example CJI format which captures the essential elements would be: Assess the ability to perform Task X by SoS configuration Y under Conditions A to achieve mission desired effect Z. For this workshop Team 6 focused on the CJI of: *Assess the ability to perform battlespace deconfliction by Current/Future Command and Control (C2) SoS under a full range of military operations in order to achieve a destroyed or neutralized Threat forces in the Joint Operations Area.*

Design of Experiment (DOE)

Team 6 used an efficient DOE approach to screen for both continuous and categorical factors that were candidate key factors in influencing the SoS effectiveness. There is an evaluate-analyze-evaluate (EAE) iteration flow in the CTM used to refine the Evaluation Strategy as shown in Figure 3. Team 6 used this EAE approach to prioritize factor importance and compared the results to the expected factor performance to see if it was consistent with experience of team subject matter experts (SMEs). This process began by defining more than forty-seven different factors with levels that were summarized in the capability crosswalk. MANA was the agent-based simulation tool available to use for data farming runs.

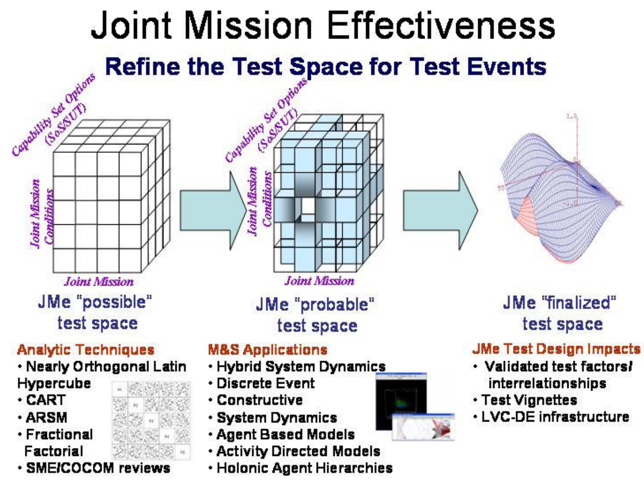

Figure 3: CTM Test Space Refinement

The capability crosswalk was mapped to a Nearly Orthogonal Latin Hypercube design and multiple iterations were run through the MANA model on the NPS cluster to provide results for further analysis and refinement of test factors. The team analyzed the responses to these runs to see if they were feasible and if expected factors were actually the most statistically significant. Throughout the week, daily replications of this process, each with numerous runs, were used to refine the test factors, based on a mission measure of effectiveness related to threat combat in a joint mission environment. The output from the initial test runs did not follow predicted factor importance. This outcome surprised many of the team members. However, further analysis of the MANA model highlighted possible limitations to the model, which were explored in subsequent DOE excursions. It was agreed that since the thrust of the week's effort was to exercise DOE processes within the CTM, the team could accept the apparent doctrinal, tactical, and performance inconsistencies in the output. The DOE process proved to be very valuable and supportive of the JTEM approach for including DOE in its methodology. Initially Team 6 had planned on running the scenario in both the MANA and Tester models. However, Tester was not available, which limited the scope of the analysis due to modeling constraints of MANA. Nonetheless, the team was able to achieve and exemplify a best practice of what could be done utilizing DOE with respect to test planning.

Analysis

Using MANA simulation results, data was processed and analyzed in order to assess SoS, threat, and environment factor importance relating to a mission measure of effectiveness. Some of the analysis outputs captured by the team are highlighted in Figures 4, 5, and 6. In Figure 4, the team applied classification and regression tree (CART) analysis output to analyze the Mission MOE concerning Proportion of Threat System Casualties. In this design, the most statistically significant factor was a forward observer (FO) "PassSelf" parameter, which turned Blue FO self reporting on or off. This PassSelf factor accounted for an R squared coefficient of determination of .587, implying that a least squares regression model relating PassSelf factor to the Mission Measure of Effectiveness (MMOE) can explain approximately 58% of the MMOE variation.

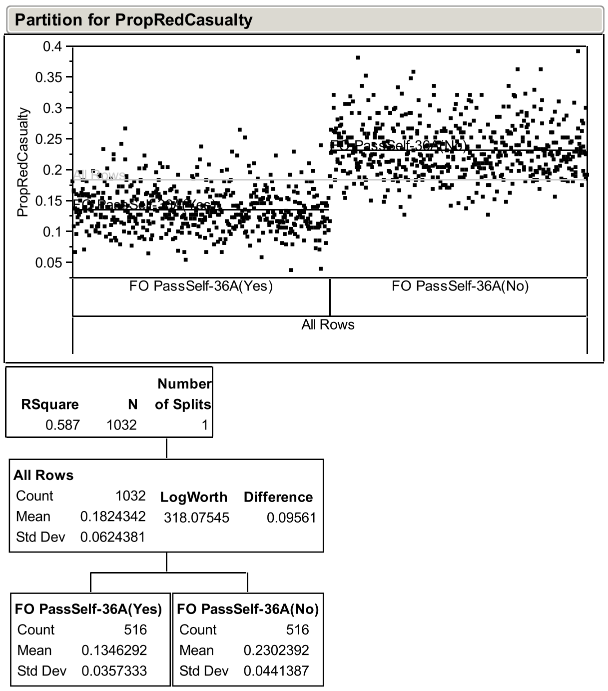

Figure 4: CART output for Red Casualties 1st Split (Left)

The partition plot in Figure 4 shows a No PassSelf parameter (no Blue FO self reporting) led to a better desired effect of threat system casualties. Similar analysis was completed to determine which variables had a smaller impact, such as the number of Red agents. Along with this analysis Pareto Plots (see Figure 5) and Prediction Profiler plots (see Figure 6) were used to model the factors that captured all main effects as well as stepwise effects chosen from all second order terms, with adjusted R squared values of .685 and .782 respectively. The adjusted R squared was essential for this analysis as it adjusts to the number of independent variables and sample size. While not as high of an R squared as preferable in smaller factor, controlled experiments, the analysis did inform the factor refinement of this large test space, which had a combination of 32 continuous and 3 categorical factors.

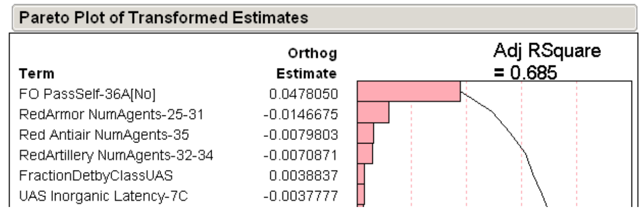

Figure 5: Pareto Plot of Model Terms of All Main Effects (Right)

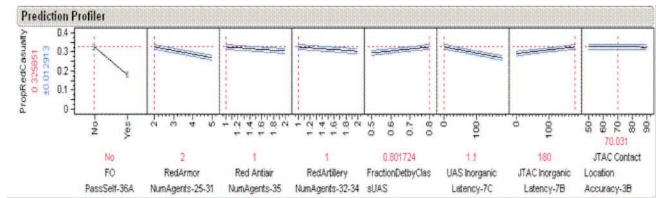

Figure 6: Profiler for 8 Most Important Factors

Insights and Issues for Further Investigation

Future data farming efforts would be more robust if they incorporated a comparison of data from both the MANA and Tester (or other models). Much of the situational awareness (SA) defined for this workshop utilizing MANA was modeled as the actual communications links both between and among entities. However the importance of C2 factors may have been underestimated. Further analysis of statistically significant C2 factors and refinement of model scenarios to better align with doctrine may provide more robust analysis. For example, if a tank battalion is down to 40% strength, and the 40% consists of the support platoon (ammunition and fuel), cooks, maintenance personnel, then that battalion has little to no fire power and would most likely not continue the offensive. The current MANA model treats many of the different elements as equal, which is not as realistic as MANA removing a unit (Blue or Red) from the fight when it reached a point of being combat ineffective. However a model is just a representation of the systems of systems "that should be used to try and gain some insights into the relationships among the various components or to predict performance under some new conditions being considered."[1]

For this analysis the attrition of Red and Blue forces was analyzed to assess the difference between the current and future SoS. Due to time constraints, this was the only MMOE or task measure of performance (TMOP) analyzed. Future investigations could expand the focus of the evaluation of this measure through an attack on the center of gravity for Red forces and then assess the impact of the loss of critical C2 nodes on Threat force attrition. However, due to the limited capabilities of Agent Based Models (ABMs), we should not throw out Red force losses and loss exchange ratios for MOEs. MOE enhancements can include weighting critical Red forces so all Red forces are not counted the same. For

[1] "Simulation, Modeling, and Analysis" Averill M. Law and W. David Kelton, p.1.

instance, the early elimination of Red air defenses could be deemed as an important desired effect achieved by joint fires and close air support tasks. Working with the MANA ABM provided an excellent opportunity to identify some needed capabilities to support MMOE and TMOP evaluations. It also provided the opportunity to identify new design aspects for the Tester model being developed. The developer for Tester was a Team 6 member and throughout the week's event he noted some critical aspects of C2 for future incorporation into future model enhancements. This immediate feedback to ABM developers is a key benefit which enhances future data farming workshops.

Having design of experiment expertise from Naval Postgraduate School, Research Development and Engineering Command (RDECOM) and TRADOC Research and Analysis Center allowed Team 6 to compare different designs to the same experimental space (NOLH, R5FF, others). This capability, combined with the application of different analytical techniques (e.g., linear regression, Kriging) allowed the team to gain valuable insights into the C2 SoS ability to perform battlespace deconfliction. Within the IDFW 16 venue, Team 6 began analyzing more than forty-eight different factors in the DOE. These factors were modeled in the MANA model and thousands of iterations were run to gain insights on the test factors and test factor interrelationships. The statistical output was then analyzed in order to validate the significance of test factors and interrelationships. This process allowed the team to indicate where models need to be changed, and where other factors or interrelationships may need to be modeled. Analysis of the data and utilization of analytic best practices such as sensitivity analysis, CART, and visualization/ analytical tools were applied to turn test data into insights including an evaluation of the overall joint mission effectiveness and the contribution a C2 SoS makes to the accomplishment of that joint mission. The IDFW 16 Workshop allowed all team members to apply a use-case focused on battlespace deconfliction to see how joint mission effectiveness test space refinement is accomplished by examining the test structures, identifying test factors and test factor interrelationships through the application of analytical techniques to identify factors of importance, factor levels of impact, and important interrelationships.